

Genotypic anomaly in Ebola virus strains circulating in Magazine Wharf area, Freetown, Sierra Leone, 2015

SL Smits¹, SD Pas¹, CB Reusken¹, BL Haagmans¹, P Pertile², C Cancedda², K Dierberg², I Wurie³, A Kamara³, D Kargbo³, SL Caddy⁴, A Arias⁴, L Thorne⁴, J Lu⁴, U Jah⁵, I Goodfellow⁴, MP Koopmans^{1,6}

1. Department of Viroscience, Erasmus Medical Center, Rotterdam, Netherlands

2. Partners in Health, Boston, United States

3. Sierra Leone Ministry of Health and Sanitation, Sierra Leone

4. University of Cambridge, Department Virology, Cambridge, United Kingdom

5. University of Makeni, School of Public Health, Sierra Leone

6. Virology Division, Centre for Infectious Diseases Research, Diagnostics and Screening, National Institute for Public Health and the Environment, Bilthoven, Netherlands

Correspondence: Saskia Smits (s.smits@erasmusmc.nl)

Citation style for this article:

Smits SL, Pas SD, Reusken CB, Haagmans BL, Pertile P, Cancedda C, Dierberg K, Wurie I, Kamara A, Kargbo D, Caddy SL, Arias A, Thorne L, Lu J, Jah U, Goodfellow I, Koopmans MP. Genotypic anomaly in Ebola virus strains circulating in Magazine Wharf area, Freetown, Sierra Leone, 2015. *Euro Surveill.* 2015;20(40):pii=30035. DOI: <http://dx.doi.org/10.2807/1560-7917.ES.2015.20.40.30035>

Article submitted on 21 September 2015 / accepted on 08 October 2015 / published on 08 October 2015

The Magazine Wharf area, Freetown, Sierra Leone was a focus of ongoing Ebola virus transmission from late June 2015. Viral genomes linked to this area contain a series of 13 T to C substitutions in a 150 base pair intergenic region downstream of viral protein 40 open reading frame, similar to the Ebolavirus/H.sapiens-wt/SLE/2014/Makona-Jo169 strain (Jo169) detected in the same town in November 2014. This suggests that recently circulating viruses from Freetown descend from a Jo169-like virus.

In Sierra Leone, two new Ebola virus (EBOV) cases were reported from the densely populated Magazine Wharf area of Freetown in the Western Area Urban district after a period of two weeks in June 2015 with no cases in the district. The Magazine Wharf area was subsequently a focus of transmission for several weeks (<http://apps.who.int/ebola/current-situation/ebola-situation-report-15-july-2015>) up to 12 August 2015 (<http://apps.who.int/ebola/current-situation/ebola-situation-report-12-august-2015>), after which no new cases were reported from the area (<http://apps.who.int/ebola/current-situation/ebola-situation-report-30-september-2015>). In this study, the whole genomes of viruses from patient samples, originating from the Western Area Urban district and other districts of the country (i.e. Kenema, Kono, and Tonkolili) between January and July 2015 are sequenced. Genomes derived from samples collected from 30 June onwards in the Western Area Urban district have a particular anomaly consisting of a series of 13 T to C substitutions in a 150 bp intergenic region downstream of the viral protein 40 (VP40) open reading frame (ORF). This anomaly is also present in a viral strain, the Ebolavirus/H.sapiens-wt/SLE/2014/Makona-Jo169 (Jo169), which was detected in Freetown in November 2014. The finding suggests

that viruses retrieved in June and July 2015 from the Western Area Urban district are direct descendants of a Jo169-like virus. Near real time application of whole EBOV genome sequencing and the identification of lineage signatures can be used to monitor the ongoing outbreak and test whether newly infected patients are part of an identified transmission chain.

Ebola virus disease epidemic in West Africa

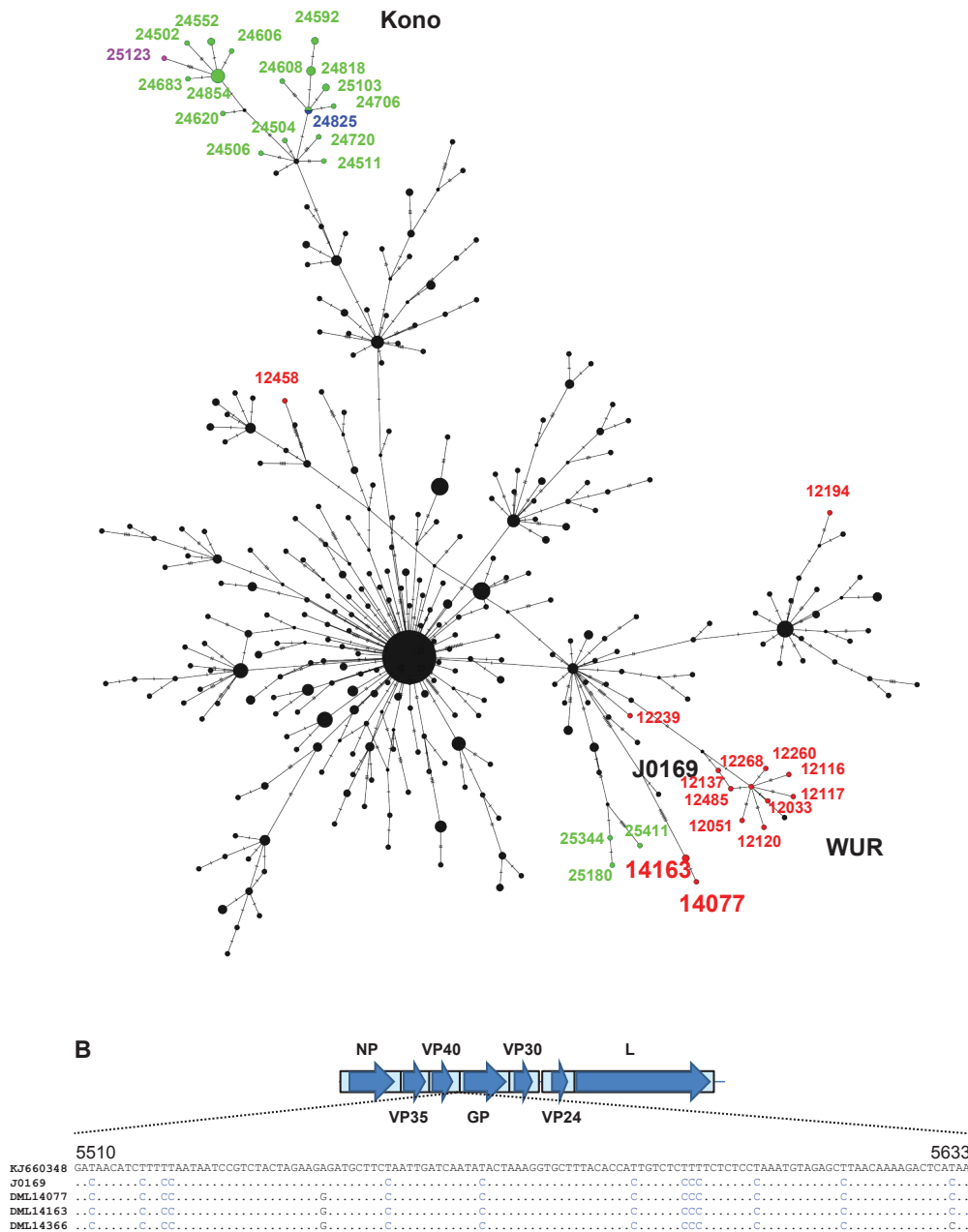
An epidemic of EBOV (a negative-sense RNA virus, family *Filoviridae*) disease has been ongoing in West Africa since December 2013 affecting mainly Guinea, Liberia and Sierra Leone [1]. As of 9 September 2015, the cumulative number of suspect, probable and confirmed cases stands at 28,183, including 11,306 deaths (<http://apps.who.int/ebola/current-situation/ebola-situation-report-9-september-2015>). EBOV cases continue to be detected and shifts in foci of transmission are observed. One of the pillars in the emergency response to the epidemic in the area has been the deployment of temporary (mobile) laboratories by the international community in collaboration with local authorities, among which the Dutch Mobile Laboratories (<http://dutchebolalabs.nl/rapport-implementatie-dutch-mobile-labs-in-sierra-leone-en-liberia/>). These laboratories provide rapid testing capacity for EBOV and malaria in support of clinical triage of (suspected) patients. In addition, the international community is currently working together in sequencing EBOV genomes in order to study EBOV evolution [2-7].

Sampling and whole genome sequencing

A total of 49 samples of EBOV positive patients who had been tested by Dutch Mobile Laboratories (Table) located in the Western Area Urban and Kono districts were included in the study. The samples were collected

FIGURE

A) Median-joining haplotype network constructed from an alignment of 563 Ebola virus sequences derived from clinical samples and B) alignment of a sequence region where four Ebola virus strains present a genotypical anomaly, Sierra Leone, 2015



GP: glycoprotein; L: RNA-dependent RNA polymerase L; NP: nucleoprotein; VP: virus protein; WUR: Western Area Urban district.

A: The 563 Ebola virus isolates' sequences correspond to nucleotides 148–18,629 of Zaire ebolavirus isolate H.sapiens-wt/SLE/2015/Makona-Goderich1 (GenBank accession number: KT345616). The median haplotype network was constructed using PopART version 1.7 (<http://popart.otago.ac.nz>). Each vertex represents a sampled viral haplotype. The size of each vertex is relative to the number of sampled isolates. Coloured vertices indicate determined viral genomes from Western Area Urban (red), Kono (green), Tonkolili (blue), and Kenema (purple) districts. The larger font numbers depicted in red represent viral genomes derived from Freetown patients' samples retrieved in June and July 2015 (two of them linked to Magazine Wharf area) with a genotypic anomaly consisting of a series of 13 T to C mutations in a 150 bp long sequence, which is located in an intergenic region downstream of the VP40 open reading frame. Hatch marks indicate single nucleotide mutations alongside each edge. The node labelled 24592 contains sequences 24592 and 24601; node 25103 contains 25083 and 25103; node 14163 contains 14163 and 14366; node 24552 contains 24552 and 24581; node 24854 contains 24854, 24853, 24605, 24669, 24677, 24604, and 24758; node 24825 contains 24825 and 24611; node 24818 contains 24818, 24553, and 24573.

B: Genotypic anomaly in four Ebola virus strains from Freetown, Western Area Urban district, consisting of a series of 13 T to C mutations in a 150 bp long sequence located in an intergenic region downstream of the VP40 open reading frame (genome positions: 5,510–5,633). Three of the strains, which are depicted as DML14077, DML14163, and DML14366 on the alignment were characterised in June and July 2015. These are respectively EBOV_DML14077_SLe_WesternUrban_20150630 (GenBank accession number: KT357860), EBOV_DML14163_SLe_WesternUrban_20150703 (GenBank accession number: KT357858), and EBOV_DML14366_SLe_WesternUrban_20150711 (GenBank accession number: KT357859). The fourth strain, which is shown as J0169 is Ebolavirus/H.sapiens-wt/SLE/2014/Makona-J0169 (GenBank accession number: KP759706) and was characterised in November 2014. The sequences of the four strains are compared with the reference Zaire ebolavirus isolate H.sapiens-wt/GIN/2014/Makona-Gueckedou-Co5, complete genome (GenBank accession number: KJ660348). The Ebola virus genome organisation is shown for reference.

TABLE

Characteristics of Ebola virus positive specimens, which were subjected to whole genome sequencing, Sierra Leone, January–July 2015 (n=49)

Specimen ID	Specimen type	Country	District ^a	Date of specimen collection	Accession number
12033	Blood	Sierra Leone	Western_Urban	19-02-2015	KT357813
12051	Blood	Sierra Leone	Western_Urban	21-02-2015	KT357814
12116	Blood	Sierra Leone	Western_Urban	26-02-2015	KT357815
12117	Blood	Sierra Leone	Western_Urban	26-02-2015	KT357816
12120	Blood	Sierra Leone	Western_Urban	27-02-2015	KT357817
12137	Blood	Sierra Leone	Western_Urban	28-02-2015	KT357819
12194	Blood	Sierra Leone	Western_Urban	04-03-2015	KT357818
12239	Swab	Sierra Leone	Western_Urban	07-03-2015	KT357820
12260	Blood	Sierra Leone	Western_Urban	09-03-2015	KT357821
12268	Blood	Sierra Leone	Western_Urban	10-03-2015	KT357822
12458	Blood	Sierra Leone	Western_Urban	28-03-2015	KT357823
12485 ^b	Swab	Sierra Leone	Western_Urban	31-03-2015	KT357824
24502 ^b	Blood	Sierra Leone	Kono	13-01-2015	KT357825
24504	Blood	Sierra Leone	Kono	13-01-2015	KT357826
24506	Blood	Sierra Leone	Kono	14-01-2015	KT357827
24511	Blood	Sierra Leone	Kono	14-01-2015	KT357828
24552	Blood	Sierra Leone	Kono	17-01-2015	KT357829
24553	Blood	Sierra Leone	Kono	17-01-2015	KT357830
24573	Swab	Sierra Leone	Kono	18-01-2015	KT357831
24581	Swab	Sierra Leone	Kono	19-01-2015	KT357832
24592*	Swab	Sierra Leone	Kono	20-01-2015	KT357833
24601*	Blood	Sierra Leone	Kono	20-01-2015	KT357834
24604	Blood	Sierra Leone	Kono	20-01-2015	KT357835
24605	Blood	Sierra Leone	Kono	20-01-2015	KT357836
24606	Blood	Sierra Leone	Kono	20-01-2015	KT357837
24608	Blood	Sierra Leone	Kono	20-01-2015	KT357838
24611	Swab	Sierra Leone	Kono	21-01-2015	KT357839
24620 ^b	Blood	Sierra Leone	Kono	21-01-2015	KT357840
24669	Blood	Sierra Leone	Kono	25-01-2015	KT357841
24677	Swab	Sierra Leone	Kono	25-01-2015	KT357842
24683**	Blood	Sierra Leone	Kono	26-01-2015	KT357843
24706	Blood	Sierra Leone	Kono	28-01-2015	KT357844
24708**c	Blood	Sierra Leone	Kono	28-01-2015	KT357845
24720 ^b	Blood	Sierra Leone	Kono	29-01-2015	KT357846
24758	Blood	Sierra Leone	Kono	30-01-2015	KT357847
24818	Swab	Sierra Leone	Kono	03-02-2015	KT357848
24825	Blood	Sierra Leone	Tonkolili	04-02-2015	KT357849
24853	Blood	Sierra Leone	Kono	06-02-2015	KT357850
24854	Blood	Sierra Leone	Kono	06-02-2015	KT357851
25083***	Blood	Sierra Leone	Kono	18-02-2015	KT357852
25103***	Swab	Sierra Leone	Kono	19-02-2015	KT357853
25123 ^b	Swab	Sierra Leone	Kenema	18-02-2015	KT357854
25180****	Blood	Sierra Leone	Kono	23-02-2015	KT357855
25344****	Blood	Sierra Leone	Kono	06-03-2015	KT357856
25411 ^b	Blood	Sierra Leone	Kono	10-03-2015	KT357857
13828	Blood	Sierra Leone	Western_Urban	22-06-2015	Not applicable
14077	Blood	Sierra Leone	Western_Urban	30-06-2015	KT357860
14163	Blood	Sierra Leone	Western_Urban	03-07-2015	KT357858
14366	Swab	Sierra Leone	Western_Urban	11-07-2015	KT357859

ID: identity; Western_Urban: Western Area Urban district.

In the column with specimen IDs, entries with the same number of asterisks indicate specimens derived from the same patient.

^a District of patient residence.

^b Less reliable consensus sequence due to low coverage regions.

^c Not shown in Figure, unreliable consensus sequence due to low coverage regions.

from patients residing in Kenema, Kono, Tonkolili and Western Area Urban districts between January and July 2015. Nucleic acids were extracted from the samples using EZ Advanced XL automated RNA extraction (Qiagen). Isolated nucleic acids were subjected to reverse transcription/polymerase chain reaction (PCR) amplification using the Ion AmpliSeq Ebola Panel Assay (Thermo Fisher Scientific) and the Ion Torrent sequencing platform at a local sequencing facility established at the Mateneh Ebola Treatment Centre in Makeni, Bombali district, Sierra Leone (European Nucleotide Archive Study: PRJEB10265). Reads were extracted from unfiltered BAM files using CLC Genomics Workbench 7.5.1 and trimmed based on quality with an ambiguous limit of 2 and quality limit of 0.05. Reads longer than 1,000 nucleotides (nt) or shorter than 15 nt were discarded. Trimmed reads were mapped to Zaire ebolavirus isolate H.sapiens-wt/GIN/2014/Makona-Gueckedou-Co5, complete genome (GenBank accession number: KJ660348), using CLC Genomics Workbench 7.5.1 map reads to reference beta module, with the following parameters: no masking, mismatch cost 2, insertion and deletion open cost 7, insertion and deletion extend cost 3, length fraction 0.95, similarity fraction 0.9. Consensus sequences were extracted and regions where depth of coverage was less than 2 were called as 'N'. All generated genomes were manually inspected for accuracy, such as for the presence of intact ORFs and low coverage regions for adequate base calling, resulting in 48 near full-length EBOV genomes (Table).

Detection of evolutionary lineages and mutations

A median haplotype network was constructed in PopART version 1.7 (<http://popart.otago.ac.nz>) using 563 determined EBOV genomes from Sierra Leone, including those determined in this study ([3-5,7]; Figure A; Table). Accordingly, the 48 determined genome sequences appeared to belong to multiple distinct evolutionary lineages. Most viruses from the Western Area Urban district grouped together as did viruses from Kono.

Viruses EBOV_DML14077_SLe_WesternUrban_20150630, EBOV_DML14163_SLe_WesternUrban_20150703, and EBOV_DML14366_SLe_WesternUrban_20150711 (DML14077, DML14163, DML14366) isolated from patients from Freetown at the end of June and July (Table) were highly similar to each other and were clearly different from viruses isolated in Freetown between January and March 2015 (Figure A). Two of these viruses were linked to the Magazine Wharf area in Freetown, a hotspot for EBOV transmission in June and July 2015.

The DML14077, DML14163, and DML14366 viruses were most closely related to Ebolavirus/H.sapiens-wt/SLE/2014/Makona-Jo169 (GenBank accession number: KP759706) isolated on 9 November 2014 in Freetown (Figure A). The Jo169 virus had a striking genotypic anomaly containing a series of 13 T to C substitutions occurring in a region of 150 bp in length in an intergenic

region downstream of the viral protein 40 (VP40) ORF. This anomaly was also present in DML14077, DML14163 and DML14366 (Figure B). An excessive accumulation of T to C mutations has been previously observed in a variety of virus genomes, including negative-sense RNA viruses and in EBOV in infected cells in vitro, and are a hallmark of host adenosine deaminases acting on RNA (ADARs) [8-10]. DML14077, DML14163, and DML14366 had 19, 17, and 18 additional single nt variants, respectively, compared with the Jo169 virus, consistent with current estimates of evolutionary rate in the Ebola virus outbreak [5], suggesting that DML14077, DML14163 and DML14366 are direct descendants from a Jo169-like virus (Figure A). The same stretch of T to C substitutions was found in a partial genome sequence from another patient at the end of June, who was also linked to the Magazine Wharf area (DML13828; Table; data not shown) and it serves as a lineage signature allowing identification and tracking of transmission chains.

Discussion and conclusions

At present only a handful of Ebola virus genomes sequenced from the current outbreak contain stretches of T to C mutations, and these can be found at different locations in the genome [5,7]. All shared nt variants, such as the conserved T to C mutations observed between Jo169 [7] and the three viral genomes newly characterised in this study, are unlikely the source of recurrent induced mutations in individual patients. Rather, the T to C substitutions could have been fixed in viral lineages in the past and these lineages may be capable of efficient human-to-human transmission. This would explain how a Jo169-like virus with a characteristic series of T to C mutations could have further spread in the Magazine Wharf area of Freetown.

It is unlikely that the T to C mutation stretch observed in the Jo169-like viruses had a fitness advantage over other spreading EBOV lineages as in the ca 6-7 months of transmission between the first identification of this virus lineage (November 2014) and the present, this EBOV lineage did not become dominant over others. The Jo169-like virus lineage seemingly lingered on at sufficiently low prevalence that it was not captured in genomic surveillance until now. This is not surprising given that the amount of whole genome sequences available is only a fraction of the 28,183 confirmed, probable and suspected cases (<http://apps.who.int/ebola/current-situation/ebola-situation-report-9-september-2015>) and whole genome sequencing started to build momentum relatively late in the outbreak. It shows however, that near real time application of whole EBOV genome sequencing and the identification of lineage signatures can be used to monitor the ongoing outbreak and test whether newly infected patients are part of an existing known transmission chain in the area.

Acknowledgements

We thank the hospital staff of PCMH, ETU, ODCH, and Koidu ETU for their efforts in the Ebola virus outbreak. We thank the Office of the President Sierra Leone, the Sierra Leone Ministry of Health and Sanitation Sierra Leone, and Partners in Health for continuous support and efforts in the Ebola virus outbreak response. We thank the drivers, pilots, non-governmental organisations, district medical and surveillance officers for their help with sample collection and logistics in Sierra Leone. We thank Bart Kooi, Adam Hume, Amber Hendriks, Anne van der Linden, Annelies Mesman, Annemieke Dinkla, Arie Kant, Bas van der Veer, Chantal Reusken, Cindy Warringa, Eline Verheij, Ellis Meuleman, Emily Nelson, Emre Kucukkose, Erik van Hannen, Evelien Kern, Felix Geeraedts, Florine Scholte, Gudrun Freidl, Guido van der Net, Hans Naus, Harry Vennema, Heidi De Gruyter, Hein Sprong, Heleen Klos, Henny Pieterse Bruins, Irma Verhoofstad, Jaco Verweij, Jacqueline Schelfaut, Janienne Klaasse, Janke Schinkel, Janko van Beek, Jelke Fros, Jennifer de Jong, Jeroen Cremer, Jeroen Roose, Jet Kant, Joffrey van Prehn, Jolanda Maaskant, Judy Yen, Jurre Siegers, Katja Wolthers, Lesla Bruijnesteijn, Lisa de Ruijter, Ludo Oostendorp, Marion Nederpel, Marion Sunter, Marloes Dullaart, Martine Slop, Matthew McCall, Nancy Beerens, Naomi Berkeveld, Nathalie van Burgel, Nicole van Manen, Nienke van Teijlingen, Nina Hertoghs, Petra van den Doel, Pieter Smit, Richard de Boer, Robbert Bentvelsen, Robin van Houdt, Ruud Jansen, Saskia Bierman, Seta Jahfari, Stefan Boers, Stephanie Blanc, Suzan Pas, and Viviana Cobos for operating the Dutch Mobile Laboratories. We thank all other people involved in getting the Dutch Mobile Laboratories up and running. We thank Terry Jones and Matt Cotton for support in transferring sequence data to the Netherlands. We thank Andrew Rambaut for helpful discussions in preparing this manuscript. This work was supported by ZonMW TOP project 91213058, the European Union's Horizon 2020 research and innovation programme under grant agreement No 643476 (COMPARE), and the Wellcome trust under grant 097997/Z/11/A.

Conflict of interest

None declared.

Authors' contributions

SLS, SDP, CBR, BLH and MPK were involved in data analysis, training of people operating Dutch Mobile Laboratories in Sierra Leone, and writing the manuscript. PP, CC, KD, IW, AK, and DK were involved in getting Dutch Mobile Laboratories operational and providing continuous support in gathering clinical samples. SLC, AA, LT, JL, UJ, and IG were involved in next generation sequencing in Sierra Leone and writing the manuscript.

References

1. BaizeS, PannetierD, OestereichL, RiegerT, KoivoguiL, MagassoubaN, et al. Emergence of Zaire Ebola virus disease in Guinea. *N Engl J Med*. 2014;371(15):1418-25. DOI: 10.1056/NEJMoa1404505 PMID: 24738640
2. CarrollMW, MatthewsDA, HiscoxJA, ElmoreMJ, PollakisG, RambautA, et al. Temporal and spatial analysis of the 2014-2015 Ebola virus outbreak in West Africa. *Nature*. 2015;524(7563):97-101. DOI: 10.1038/nature14594 PMID: 26083749
3. GireSK, GobaA, AndersenKG, SealfonRS, ParkDJ, KannehL, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science*. 2014;345(6202):1369-72. DOI: 10.1126/science.1259657 PMID: 25214632
4. US Army Medical Research Institute of Infectious Diseases, KugelmanJR, WileyMR, MateS, LadnerJT, BeitzelB, FakoliL, et al. Monitoring of Ebola Virus Makona Evolution through Establishment of Advanced Genomic Capability in Liberia. *Emerg Infect Dis*. 2015;21(7):1135-43. DOI: 10.3201/eid2107.150522 PMID: 26079255
5. ParkDJ, DudasG, WohlS, GobaA, WhitmerSL, AndersenKG, et al. Ebola Virus Epidemiology, Transmission, and Evolution during Seven Months in Sierra Leone. *Cell*. 2015;161(7):1516-26. DOI: 10.1016/j.cell.2015.06.007 PMID: 26091036
6. Simon-LoriereE, FayeO, FayeO, KoivoguiL, MagassoubaN, KeitaS, et al. Distinct lineages of Ebola virus in Guinea during the 2014 West African epidemic. *Nature*. 2015;524(7563):102-4. DOI: 10.1038/nature14612 PMID: 26106863
7. China Mobile Laboratory Testing Team in Sierra Leone, TongYG, ShiWF, LiuD, QianJ, LiangL, BoXC, et al. Genetic diversity and evolutionary dynamics of Ebola virus in Sierra Leone. *Nature*. 2015;524(7563):93-6. DOI: 10.1038/nature14490 PMID: 25970247
8. SamuelCE. Adenosine deaminases acting on RNA (ADARs) are both antiviral and proviral. *Virology*. 2011;411(2):180-93. DOI: 10.1016/j.virol.2010.12.004 PMID: 21211811
9. CattaneoR, SchmidA, EschleD, BaczkokK, ter MeulenV, BilleterMA. Biased hypermutation and other genetic changes in defective measles viruses in human brain infections. *Cell*. 1988;55(2):255-65. DOI: 10.1016/0092-8674(88)90048-7 PMID: 3167982
10. ShabmanRS, JabadoOJ, MireCE, StockwellTB, EdwardsM, MahajanM, et al. Deep sequencing identifies noncanonical editing of Ebola and Marburg virus RNAs in infected cells. *MBio*. 2014;5(6):e02011. DOI: 10.1128/mBio.02011-14 PMID: 25370495